REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
25/07/2014

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
01/09/2010 - 31/08/2012

4. TITLE AND SUBTITLE
BAYESIAN TRACKING WITHIN A FEEDBACK SENSING ENVIRONMENT: Estimating Interacting, Spatially Constrained Complex Dynamical Systems from Multiple Sources of Controllable Devices

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-10-1-0501

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Emily B. Fox

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Wharton School, University of Pennsylvania
400 Jon M. Huntsman Hall
3730 Walnut Street
Philadelphia, PA 19104-6340

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Science and Research
875 Randolph Street
Suite 325 Room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This grant led to developments in flexible models for complex time series in a range of applications with a focus on Bayesian and Bayesian nonparametric methods. Three fundamental challenges were tackled: (i) capturing evolving correlations in high-dimensional time series with possible missing or irregularly-spaced observations, (ii) performing diverse subset selection over time, and (iii) automatically learning an unknown set of simple underlying temporal structures to describe complex dynamical phenomena. Each of these methods was applied in a range of application domains including neuroimaging, diverse document selection, speaker diarization, stock modeling, and target tracking.

15. SUBJECT TERMS
Bayesian modeling, Bayesian nonparametrics, dynamical systems, Gaussian processes, Markov switching processes, latent dict

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	UU	4	Emily B. Fox
					19b. TELEPHONE NUMBER *(include area code)* 206-221-9341

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Subject: Final Report to Dr. Tristan Nguyen

Contract/Grant Title:	Bayesian Tracking within a Feedback Sensing Environment: Estimating Interacting, Spatially Constrained Complex Dynamical Systems from Multiple Sources of Controllable Devices
Contract/Grant #:	FA9550-10-1-0501
Principal Investigator:	Emily Fox
Reporting Period:	1 September 2010 to 31 August 2012
Date:	25 July 2014

Abstract:

This grant led to developments in flexible models for complex time series in a range of applications with a focus on Bayesian and Bayesian nonparametric methods. Three fundamental challenges were tackled: (i) capturing evolving correlations in high-dimensional time series with possible missing or irregularly-spaced observations, (ii) performing diverse subset selection over time, and (iii) automatically learning an unknown set of simple underlying temporal structures to describe complex dynamical phenomena. We applied the first of these techniques to a task of classifying word stimuli based on 102-dimensional MEG neuronal time-series responses, and achieved state-of-the-art performance improving upon an SVM classifier. By modeling changing correlations, we are also able to infer cortical regions (i.e., groups of response trajectories) with coordinated activity. For the diversity modeling approach, we considered a task of selecting diverse yet relevant news articles to display to users over time. For a simulated user with unknown preferences over topics, our method had better precision/recall performance than competing methods, more rapidly discovering the articles preferred by the user. Finally, we applied the model based on a composition of simple temporal structures to a speaker diarization task with the goal of segmenting conference audio in the presence of an unknown number of speakers. Our Bayesian nonparametric approach outperformed a highly-engineered gold-standard method on the standard NIST dataset. Building on the same model, we were also able to segment dances of honey bees, volatility in the IBOVESPA stock index, and formulate a target tracking application.

Final accomplishments:

The methods developed under this grant are based on three fundamentally different approaches to modeling complex time series. The first, as considered in [1], learns a latent dictionary of Gaussian processes to model high-dimensional time series with possibly missing and irregularly spaced observations. A key feature of this model is the ability to capture continually changing correlations between the many dimensions of the observation vector over time. The second, as considered in [2], employs a separate type of random process: a determinantal point process (DPP), which is a repulsive process useful in diverse subset modeling. We developed a time series version of this random process that captures diversity of the subset at each time step as well as diversity of subsets between time steps. Finally, in [3]-[5] we further developed and extended the work on Bayesian nonparametric switching linear dynamical systems outlined as preliminary work in the original proposal. Each of these methods was applied in a range of application domains including neuroimaging, diverse document selection, speaker diarization, stock modeling, and target tracking. We detail each of these projects further below. The impact on the community can be summarized as developing methods to:

- capture evolving correlations in high-dimensional time series with possible missing or irregularly-spaced observations
- performing diverse subset selection over time
- automatically learning an unknown set of simple underlying temporal structures to describe complex dynamical phenomena.

In terms of application domains, the impact was:

- a state-of-the-art method for classifying word stimuli based on single-trial MEG neuronal recordings
- ability to infer cortical regions (i.e., groups of MEG response trajectories) with coordinated activity
- a technique for displaying diverse and high-quality news articles to a user, with better precision/recall performance than competing methods in a task of discovering articles preferred by the user
- a gold-standard speaker diarization method, as demonstrated on the standard NIST dataset
- a model able to segment dances of honey bees or volatility in the IBOVESPA stock index.

Latent Dictionary Learning for High-Dimensional Evolving Correlations In [1], we proposed a hierarchical latent dictionary approach to estimate the time-varying mean and covariance of a high-dimensional process for which we have only limited noisy samples. Most previous methods have focused just on capturing a time-varying mean. However, in many application domains it is insufficient to assume that the correlations between the elements of the observation vector are static. For example, the spatial correlation of Magnetoencephalography (MEG) sensors change as the co-activation pattern of brain regions evolves in time. In such cases, one needs a heteroscedastic model. A challenge is both the dimensionality of the time series resulting (e.g., many MEG sensors) as well as the fact that the signal to noise ratio can be extremely low given the recording setup (e.g., non-invasive cortical readings).

To cope with the dimensionality of the time series, we leverage the typical redundancy in sensor measurements by considering lower dimensional latent processes. To capture potential long-range dependencies, we take the latent trajectories, or dictionary elements, to be Gaussian process random functions. This formulation also enables us to cope with possible missing values (per time step) or an irregular grid of observations, both of which are useful in many real world applications. In scenarios where multiple related recordings are collected, as in the MEG application, we devised a variant of the model that hierarchically couples the latent trajectories to transfer knowledge between the multiple recordings and better recover the signal from few noisy samples.

In [1], we apply our methods to the task of identify the word being viewed by a human subject based solely on MEG recordings of their brain activity. Specifically, we identify the word category for a single noisy MEG recording, when only given limited noisy samples on which to train. Our model provides the current gold-standard for this task, outperforming powerful discriminative methods (e.g., SVM classifiers) and additionally affords many opportunity for extensions being based on a generative model.

Diverse Subset Modeling Over Time A determinantal point process (DPP) is a random process useful for modeling the combinatorial problem of subset selection. In particular, DPPs encourage a random subset Y to contain a diverse set of items selected from a base set $\mathcal{Y}$. For example, we might use a DPP to display a set of news headlines that are relevant to a users interests while covering a variety of topics. Suppose, however, that we are asked to sequentially select multiple diverse sets of items, for example, displaying new headlines day-by-day. We might want these sets to be diverse not just individually but also through time, offering headlines today that are unlike the ones shown yesterday. In [2], we constructed a Markov DPP (M-DPP) that models a sequence of random sets $\{Y_t\}$. The proposed M-DPP defines a stationary process that maintains DPP margins. Crucially, the induced union process $Z_t = Y_t \cup Y_{t1}$ is also marginally DPP-distributed. Jointly, these properties imply that the sequence of random sets are encouraged to be diverse both at a given time step as well as across time steps. We derived an exact, efficient sampling procedure, and a method for incrementally learning a quality measure over items in the base set $\mathcal{Y}$ based on external preferences.

We applied the M-DPP to the task of sequentially displaying diverse and relevant news articles to a user with topic preferences. We found empirically that the model achieves an improved balance between diversity and quality compared to baseline methods. We also studied the effects of the M-DPP on learning, finding significant improvements in recall at minimal cost to precision for a news task where user feedback was provided.

Bayesian Nonparametric Learning of Markov-Switching Dynamical Systems Markov switching processes, such as the hidden Markov model (HMM) and switching linear dynamical system (SLDS), are often used to describe rich dynamical phenomena. They describe complex behavior via repeated returns to a set of simpler models: imagine a person alternating between walking, running, and jumping behaviors, or a stock index switching between regimes of high and low volatility. Classical approaches to identification and estimation of these models assume a fixed, pre-specified number of dynamical models. In [3]-[5], we instead examined Bayesian nonparametric approaches that define a prior on Markov switching processes with an unbounded number of potential model parameters (i.e., Markov modes). By leveraging stochastic processes such as the Dirichlet process, these methods allow the data to drive the complexity of the learned model, while still permitting efficient inference algorithms. One key contribution of this work was formulating a process that captures the natural persistence of dynamical modes present in many real world processes. We referred to this model as the *sticky HDP-HMM* (hierarchical Dirichlet process hidden Markov model).

In [3], we applied the sticky HDP-HMM to the problem of speaker diarization where the goal is to segment an audio recording of a meeting into temporal segments corresponding to individual speakers. The problem is rendered particularly difficult by the fact that we are not allowed to assume knowledge of the number of people participating in the meeting. Although the basic HDP-HMM tends to over-segment the audio data–creating redundant states and rapidly switching among them–the sticky HDP-HMM provides effective control over the switching rate by capturing in the prior that if you are currently speaking you are more likely to continue speaking than to transition to a new speaker (i.e., temporal persistence of states). We also show that this sticky HDP-HMM makes it possible to treat the observation model (emissions) nonparametrically. Accommodating multimodal emissions is essential for the speaker diarization problem and is likely an important ingredient in other applications of the HDP-HMM. Working with a benchmark NIST data set, we showed that our Bayesian nonparametric architecture yields state-of-the-art speaker diarization results.

Finally, to scale the resulting architecture to realistic diarization problems, we developed a Markov chain Monte Carlo (MCMC) sampling algorithm that employs a truncated approximation of the Dirichlet process to jointly resample the full state sequence using a variant of the forward-backward algorithm, greatly improving mixing rates.

In [4]-[5], we considered more complex state-specific dynamical models that extend the sticky HDP-HMM's assumption of conditionally independent observations over time (conditioned up the state sequence). In particular, we consider two such systems that switch among a set of conditionally linear dynamical modes: the switching linear dynamical system (SLDS) and the switching vector autoregressive (VAR) process. We additionally employ a sparsity inducing prior—automatic relevance determination—to infer a sparse set of dynamic dependencies allowing us to learn SLDS with varying latent state dimension or switching VAR processes with varying autoregressive order. We developed an MCMC sampling algorithm that combines a truncated approximation to the Dirichlet process with efficient joint sampling of the mode and state sequences. We demonstrated the utility and flexibility of our model on segmenting sequences of dancing honey bees, the IBOVESPA stock index, and in a maneuvering target tracking application. In each of these significantly different application domains, we use the same fundamental modeling building block and parameter settings. In one case we are able to learn changes in the volatility of the IBOVESPA stock exchange while in another case we learn segmentations of data into waggle, turn-right, and turn-left honey bee dances. These results illustrate the importance of our model's ability to automatically discovering simple underlying temporal structures to describe the complex dynamical phenomena.

Archival publications (published) during reporting period:

- [1] A.M. Fyshe, E.B. Fox, D.B. Dunson, and T.M. Mitchell, "Hierarchical Latent Dictionary Learning for Word Classification using Brain Activation Patterns," Proc. Intl. Conf. on Artificial Intelligence and Statistics, JMLR 22: 409-421, 2012.

- [2] R.H. Affandi, A. Kuleza, and E.B. Fox, "Markov Determinantal Point Processes," Proc. Uncertainty in Artificial Intelligence, 2012.

- [3] E.B. Fox, E.B. Sudderth, M.I. Jordan, and A.S. Willsky,"A Sticky HDP-HMM with Application to Speaker Diarization," Annals of Applied Statistics, vol. 5, no. 2A, pp. 1020-1056, June 2011.

- [4] E.B. Fox, E.B. Sudderth, M.I. Jordan, and A.S. Willsky, “Bayesian Nonparametric Inference of Switching Dynamic Linear Models,” IEEE Transactions on Signal Processing, vol. 59, no. 4, pp. 1569-1585, April 2011.
- [5] E.B. Fox, E.B. Sudderth, M.I. Jordan, and A.S. Willsky, “Bayesian Nonparametric Learning of Markov Switching Processes,” IEEE Signal Processing Magazine, vol. 27, no. 6, pp. 43-54, November 2010.

Changes in research objectives, if any: None

Change in AFOSR program manager, if any: None

Extensions granted or milestones slipped, if any: None

Include any new discoveries, inventions, or patent disclosures during this reporting period (if none, report none): None